February 21, 2007

A Review of the December 2006 University of California-Davis Report entitled "Methodology for Derivation of Pesticide Water Quality Criteria for the Protection of Aquatic Life in the Sacramento and San Joaquin River Basins: Phase II: Methodology Development and Derivation of Chlorpyrifos Criteria" by Patti TenBrook and Ronald Tjeerdema

By

Lenwood Hall, Jr.
University of Maryland
Wye Research and Education Center
P. O. Box 169
Queenstown, Maryland 21658

The goal of this project is to develop a new methodology for derivation of pesticide water quality criteria for the protection of aquatic life in the Sacramento and San Joaquin River Basins. This project is funded by California's Central Valley Regional Water Quality Control Board. The project has three phases. Phase I was a global comparison and analysis of existing water quality criteria methodologies published in April of 2006 (TenBrook and Tjeerdema, 2006). I provided a previous review of this phase I report in June of 2006 (Hall, 2006). The Phase II report, which is the subject of this review, is a presentation of the new criteria development methodology. This phase II report also includes a chapter that uses the new methodology to derive acute and chronic criteria for chlorpyrifos. A phase III report will be produced later in 2007 that applies the new methodology to derive criteria for up to five pesticides (including diazinon). General comments on this phase II report are included below followed by specific page by page comments.

General Comments

- **Specific Transparent Objectives** - A clear statement of specific goals is needed in the Introduction. The current text states that the goal is to "develop a methodology for derivation of pesticide water quality criteria for the protection of aquatic life in the Sacramento and San Joaquin River Basins". The critical point in this goal is what level of protection does the new method seek, i.e. protection of all species, 95% of the species as outlined in the USEPA water quality criteria document (Stephen et al. 1985) or some other level of protection. USEPA assumes that aquatic ecosystems can tolerate some stress and occasional adverse effects; therefore, protection of all species at all times and places is not necessary (Stephen et al., 1985). It is also unclear if the new methodology would apply to pesticides such as copper (a trace metal). If so, then the water quality effects section would need to be expanded to address water quality effects (i.e., hardness and dissolved organic carbon influence copper toxicity). The Introduction should

also state why the Central Valley Regional Water Quality Control Board (CVRWQCB) has decided that a new criteria derivation method is needed. Does the CVRWQCB believe the existing criteria development methods used by USEPA and California Department of Fish and Game (CDFG) are inadequate or in some way flawed?

- **Data Driven Process** - The new criteria development methodology should be "data driven" and require at least as much toxicity data as the USEPA method (Stephen et al., 1985) to avoid uncertainty in the final acute and chronic criteria for pesticides. However, this is not the case as data for only 5 species are required with this new methodology for an SSD compared to 8 species required by USEPA (Stephen et al. 1985). The use of 5 toxicity data points is problematic as Wheeler et al. (2002) states that 10 toxicity data points from individual species are needed for a reasonable SSD. An even more troubling component of the new methodology is the use of assessment factors (also called safety factors, application factors and extrapolation factors) for pesticides with small data sets (less than five toxicity values for designated species). The use of assessment factors greatly increases the possibility of overestimating risk as reported by Chapman et al. (1998) and discussed by the authors. For example, the authors provide an example of how conservative the assessment factor approach can be with the chlorpyrifos example in Chapter 4. The final acute criterion derived by the new criteria method is 11.5 ng/L based on five acute data points but if only one data point had been available for *Daphnia*, the assessment factor approach would have derived an acute criterion on 0.03 ng/L. This is an extremely low value, below 1 ng/L, that can not be measured with current analytical methods. An example of unnecessary data reduction in the proposed methodology is the use of only North American species for criteria development. Given that the presence of limited toxicity data is a major issue with criteria development it would seem prudent to use a phylogenetic rather than geographic considerations when selecting toxicity data. If toxicity data were available for a non native North American species that has a closely related species present in North America then these data should be used for criteria development if the study is acceptable based on the data screening process. This approach would be acceptable since Suter 1993 has demonstrated that closely related species have similar sensitivity to contaminants. As stated by the authors, the best way to minimize overprotection and provide science based criteria is to expand available acute and chronic toxicity data sets. I would strongly support this recommendation and promote a science based "data driven" approach.
- **Policy Decisions** -There is a continual theme throughout the report that various critical components of the new methodology are policy decisions, i.e. acceptance of certain toxicity data, selection of certainty levels in tails of species sensitivity distributions, and determination of assessment factors. I strongly disagree with this approach because empirical science should be used to determine the various critical components of the new criteria methodology. Both qualified scientists and policy types should work together to develop the various components of the new criteria methodology.

- **Averaging Periods** – For acute criteria, a 1-h averaging is proposed while for chronic criteria a 4-d period is established. These two averaging periods are used by USEPA in their criteria development method (Stephen et al., 1985). It is important to remember that the USEPA approach developed in the mid 1980s was primarily developed for POINT SOURCE discharges where constituents such as ammonia are measured at frequent intervals (hourly or daily). However, for pesticides hourly measurements are rare for monitoring effects in California. Daily measurements for four consecutive days are somewhat more likely but are still the exception and not the rule for pesticide monitoring studies in the Central Valley. Therefore, the basis for using 1-h (acute criterion) and 4-d (chronic criterion) averaging periods for allowable exposure duration for pesticides in the Central Valley is not appropriate. Pesticide data collected from monitoring studies in the Central Valley and obtained from California's Department of Pesticide Regulation should be reviewed to determine the most common frequency of pesticide measurements (i.e., once a month for a year) and these data should be used to select the most appropriate averaging periods for both acute and chronic criteria.
- **Frequency of Exceedence** – In setting an allowable frequency of exceedence of acute and chronic criterion, the key question is how much time is needed for organisms at various levels of organization to recover from brief pulse exposures to contaminants. The proposed criteria method recommends an allowable frequency of exceedance of once in three years. This is the same frequency of exceedance used by the USEPA in their criteria method (Stephen et al., 1985) and as stated by the authors the 3-year frequency of exceedence was supported by minimal data. The receptor group (most sensitive biological assemblage) for any given pesticide should be considered when establishing the frequency of exceedance for a specific type of pesticide. For example, the receptor group for herbicides is plants such as phytoplankton which have short life histories (several days). Therefore, a once in three years exceedance is overprotective for species such as phytoplankton which can recover within days or weeks. In contrast, for species with long life cycles (greater than 5 years) such as various fish, a once in three year exceedance may be appropriate. I would recommend flexibility for the frequency of exceedance component of the new criteria development method that would allow the use of life histories for receptor species in order to determine the most appropriate frequency of exceedance. The authors should also explore the use of the binomial approach for determining the number of pesticide exceedences needed before a violation occurs. The California State Board uses the binomial approach for listing and delisting impaired water bodies in the State based on exceedences of both toxicants (i.e. pesticides) and conventional pollutants (i.e., pH, dissolved oxygen) (SWRCB, 2004). The binomial approach has statistical underpinnings that allows the determination of error rates associated with impairment declarations and a process to limit error rates.
- **Justification of 5th percentile -** Using a species sensitivity distribution (SSD) for criteria derivation requires selection of a percentile of the distribution as a cutoff point. An interpretation of this cutoff point means that species lying above this point in the distribution will be protected as long as the concentration of the

chemical is below the concentration at the selected percentile. The authors state that species lying below percentile would be harmed. This is incorrect. Species lying below this percentile would not be fully protected but not necessarily harmed. The authors state that the choice of the 5th percentile is purely pragmatic and has been used by other organizations such as USEPA (Stephen et al., 1985), the Dutch (RIVM, 2001), and Australia/New Zealand (ANZECC and ARMCANZ, 2000) without rigorous scientific justification. Therefore, scientific rationale should be provided before the 5th percentile is used in the new criteria development method. In addition, scientific rationale should be provided to justify dividing the 5th percentile by a factor of 2 before determining the final acute value.

- **Harmonization/coherence across media –** As stated by the authors, the final element to consider is whether a pesticide that is present in water at a criterion level might have the potential to move from that water compartment into another environmental compartment (i.e., sediment, biota, air). This harmonization issue will be particularly important for hydrophobic pesticides, such as pyrethroids, that may eventually concentrate in bed sediment. Therefore, water water quality criteria and sediment criteria for pesticides such as pyrethroids need harmonization to avoid possible conflicts. This would involve communication between Regional Board scientists and State Board scientists (i.e., Chris Beegan) that are addressing these water quality and sediment quality criteria issues.

Specific Comments

Specific comments by page are listed below.

Page ii, paragraph 1 – When will the Phase III report be available? What process will follow the Phase III report? Will a basin plan amendment be developed to approve the new criteria methodology? Will documentation be provided on how the Regional Board has responded to review comments from the public/interested parties?

Page iii, paragraph 2 – I would suggest rounding off the acute chlorpyrifos criterion to 12 ng/L and the chronic criterion to 11 ng/L to reflect the sensitivity of the analytical method for chlorpyrifos measurements. It is also prudent to check monitoring data for chlorpyrifos in the Central Valley to see if the current analytical detection limits for chlorpyrifos used by most laboratories are below or above the proposed criteria.

Page 2-1, paragraph 1 – The last sentence in this paragraph states that 11 other pesticide data sets were used from EPA; however, only 9 references are provided.

Page 2-2, Acute – Acute methodologies for plants should be included.

Page 2-6, paragraph 3 – It is unclear how the use of non-traditional endpoints may be used to derive criteria if those endpoints have been adequately linked to effects on survival, growth and reproduction or population parameters. Who makes this very

critical decision on the use of non-traditional endpoints for criteria derivation? (a panel of experts, Regional Board scientists).

Page 2-6, Multispecies – Microcosm and mesocosm data should used in the criteria derivation process if it is available and valid. For example, if microcosm/mesocosm NOECs/LOECs are substantially higher that the acute or chronic criterion then the data used to develop the criteria should be reevaluated.

Page 2-14, paragraph 1 – The rationale behind using the 75th percentile of scores for the reliability rating is needed. Chlorpyrifos may not be a good data set to use for this benchmark since this is a fairly rich data set and most other pesticide toxicity data sets will be less extensive.

Page 2-17, paragraph 4 – The point concerning considerable variability of sensitivity between species within a genus is generally not supported by most of the literature. Suter (1993) that showed similar species have similar response to chemicals.

Page 2-21, table 2.1 – Why were these 12 pesticide selected as test cases for the SSD method? Do they cover all the classes of pesticides (i.e. organophosphates, herbicides) that are used in the Central valley?

Page 2-37, paragraph 2 – Rather than prescribe the distribution to use for the pesticide toxicity data, Burr III distribution, why not use the distribution that best fits the data?

Page 2-47 and 2-48 – The points made by Chapman et al. (1998) and discussed by the authors would seem to justify why "Assessment Factors" should not be used to establish criteria, i.e. it greatly increases the possibility of overestimating risk (see my previous comments in general comments section of this review).

Page 2-50, point #1 – The use of toxicity data from the daphnia family for limited toxicity data sets may be overprotective for OP insecticides, since the receptor and most sensitive taxa are cladocerans. However, the use of daphnids may be underprotective or inappropriate for other pesticides where daphnids are not the receptor taxa.

Page 2-51, Table 2.6 – This table provides a clear example of why limited toxicity data ($n<5$) should never be used to establish criteria.

Page 2-52, paragraph 3 – Saltwater taxa should only be used if the pesticide toxicity data shows that salinity does not affect the toxicity of the pesticide. For example, salinity affects the toxicity of metals such as copper.

Page 2-54, paragraph 1 and Table 2.8 – Why did the Great Lakes guidance document select the 80th percentile as a default value of ACRs? It should be stated clearly in this paragraph if an ACR is available for a pesticide (as is the case for chlorpyrifos) then this ACR is used and not the default value of 12.4. The ACR for lindane is higher

than the other ACRs in Table 2.8 and should therefore be checked carefully before including in this table

Page 2-54, Averaging Periods – It would be more reader friendly to include the frequency and duration components within the same section or subsection since these components of the criteria are closely tied together.

Page 56, paragraph 2 – The comments that chlorpyrifos and diazinon are not fast acting toxicants is not supported by the newly derived chlorpyrifos acute and chronic values which are nearly identical and a previously published EPA diazinon criterion of 100 ng/L for both the acute and chronic criteria (USEPA, 2000).

Page 2-64, paragraph 3 – Semi-permeable membrane devices (SPMDs) are passive sampling devices that are intended to mimic uptake of bioavailable contaminants. The various negatives associated with using SPMDs are presented by the authors (i.e., they do not give quantitative results for polar organics). Therefore, I would not support the use of SPMDs for assessing bioavailablity. However, if tissue data were available for resident bivalves in a particular study area potentially impacted by pesticides or well designed caging experiments with bivalves were conducted in the study areas, these data may be useful for addressing bioavailability issues. In order to address the issue of bioavailability, the new criteria method needs to have some flexibility to address this issue on a "pesticide specific basis" depending on the data available and the physical/chemical properties of the pesticide.

Page 2-67, paragaraph 2 and 3 – The authors must be careful when evaluating possible additivity of chemicals with similar modes of action. The first consideration is that the chemicals must co-occur in the environment (present in the same sample). The next consideration is that additivity can not be assumed if measured concentrations of pesticides are below a certain threshold of response (Dr. Allan Felscot, personal communication, Washington State University).

Page 2-73, paragraph 2 – It seems a stretch to include protection of terrestrial wildlife or human health within this report since the goal is to develop a criteria method for protection of aquatic life.

Page 2-75, Threatened and endangered species – It would seem appropriate to have qualified individuals from EPA, U. S. Fish and Wildlife Service and NOAA ,that work with threatened and endangered species, review this part of the criteria development method.

Pages 3-1 and 3-2 – For both the acute and chronic sections definitions/references for plant toxicity testing should be included.

Page 3-4 and 3-5 – How does the toxicity data screening process developed by the authors compare with the data screening process used by USEPA for their development of water quality criteria.

Page 3-6, last paragraph – The authors support the use of a statistical test for outliers (Sokal and Rohlf, 1995) to delete suspect data points. Although this statistical approach is admirable, the authors may also want to consider the approach used by EPA for addressing outliers. EPA has addressed this issue in Stephen et al, (1995) by stating the following "acute values that appear to be questionable in comparison to other acute or chronic values for the same species or other species in the same genus should not be used in the calculation of species mean acute values".

Page 3-10, paragraph 1 – The authors state that the BurrliOZ software comes with a caution that for data sets of 8 or fewer values there is a great uncertainty in the calculated toxicity values. This provides further support for one of my main comments (see general comments section) that the use of 5 data points is too data restrictive and will produce SSDs and 5th centiles with a high degree of uncertainty.

Page 3-B8, Table 3.7 – The first parameter listed in Table 3.7 is "results published or in signed, dated form". It is unclear to me what the term "published" means. If this means published in the "peer reviewed" literature then data reported from this type of reference should score higher than a published report that has not been subjected to "peer review".

Page 3-B16, Table 3.16 – In developing this table for ACRs that results in an ACR default value, I would suggest that ACRs for all classes of pesticides (i.e., herbicides, carbamates) that are suspected as potential stressors in the Central Valley be included in this table. This would provide a more representative ACR default value for the geographic area of concern.

Page 4-5, paragraph 4 – The chronic value of 1 ng/L for *Neomysis mercedis* is extremely low and suspect. The study and ACE analysis that derived this value needs to be carefully reviewed. The authors later mention on page 4-6 that the *Neomysis* chronic value is not used to calculate the ACR. This point should be stated on page 4-5.

Page 4-6, top of page – For transparency, the authors need to provide more details on how the 5th and 1th percentiles were determined.

Page 4-7, Mixtures, paragraph 2 – The authors state that Table 4.9 shows synergistic ratios. The Table 4.9 in my downloaded copy is *Neomysis* raw acute data from CDFG.

Page 4-14, paragraph 2 – The authors explain the differences in their new lower acute and chronic criterion for chlorpyrifos compared with the EPA values or CDFG values by stating that different data sets were used for final calculations. A table or series of tables should be developed to clearly show why the chlorpyrifos criteria are different among the three methods, i.e., the new method, the EPA method and the CDFG method.

References

ANZECC and ARMCANZ. 2000. Australian and New Zealand guidelines for fresh and marine water quality. Report. Australian and New Zealand Environment and Conservation Council and Agricultural and Resource Management Council of Australia and New Zealand, Canberra, Australia.

Chapman, P.M., A. Fairbrother, and D. Brown. 1998. A critical evaluation of safety (uncertainty) factors for ecological risk assessment. Environ. Toxicol. Chem. 17:99-108.

Hall, L.W. Jr. 2006. A review of the April 2006 University of California-Davis Report entitled "Methodology for Derivation of Pesticide Water Quality Criteria for Protection of Aquatic Life in the Sacramento and San Joaquin Basins Phase I: Review of Existing Methodologies". University of Maryland, Wye Research and Education Center, Queenstown, Maryland.

RIVM. 2001. Guidance document on deriving environmental risk limits in The Netherlands. Report. National Institute of Public Health and the Environment

Sokal, R.R. and F.J. Rohlf. 1995. Biometry, the Principles and Practice of Statistics in Biological Research, W. H. Freeman and Company, New York, New York.

State Water Resources Control Board (SWRCB). 2004. Water quality control policy for developing California's Clean Water Act Section 303 (d) List. California Environmental Protection Agency, Sacramento, California.

Stephen, C.E., D.I. Mount, D.J. Hansen, J. H. Gentile, G.A. Chapman, and W. D. Brungs. 1985. Guidelines for deriving numerical national water quality criteria for the protection of aquatic organisms and their uses. NTIS PB85-227049. USEPA Research Laboratory, Duluth, Minnesota.

Suter, G. 1993. Part III Effects Assessment. In G. Suter (eds.) Ecological Risk Assessment, Lewis Publishers, Boca Raton, Florida, pp 175-246.

TenBrook, P.L. and R.S. Tjeerdema. 2006. Methodology for Derivation of Pesticide Water Quality Criteria for the Protection of Aquatic Life in the Sacramento and San Joaquin Basins Phase I: Review of Existing Methodologies. Prepared by the University of California at Davis for the Central Valley Regional Water Quality Control Board, Rancho Cordova, California.

United States Environmental Protection Agency (USEPA). 2000. Draft Ambient Aquatic Life Water Quality Criteria Diazinon. Report, Contract Number 68-C—98-134, USEPA, Duluth, Minnesota.

Wheeler, J.R., E.P.M.Grist, K.M.Y. Leung, D. Morritt, and M. Crane. 2002. Species sensitivity distributions: data and model choices. Marine Pollution Bulletin 45: 192-202.